Cell, Vol. 51, 435–443, November 6, 1987, Copyright © 1987 by Cell Press

Genomic Footprinting Reveals Cell Type–Specific DNA Binding of Ubiquitous Factors

Peter B. Becker, Siegfried Ruppert, and Günther Schütz
Institute of Cell and Tumor Biology
German Cancer Research Center
Im Neuenheimer Feld 280
6900 Heidelberg, Federal Republic of Germany

Summary

Using in vivo dimethylsulfate footprinting, we have analyzed protein–DNA interactions within two regions upstream of the tyrosine aminotransferase (TAT) gene that are characterized by an altered chromatin structure in TAT-expressing as compared to nonexpressing cells. All the identified protein contacts to DNA are found exclusively in the TAT-expressing hepatoma cells. In vitro analyses of specific DNA-binding factors in crude nuclear extracts yield DNAase I footprints that correlate well with the binding sites in vivo. Surprisingly, all DNA-binding activities are present in nuclei of TAT-expressing and nonexpressing cells, indicating that the mere presence of factors is not sufficient for their interaction with a binding site in vivo. Genomic sequencing reveals methylation of CpG dinucleotides in the regions analyzed in nonexpressing cells, whereas no methylation is found in TAT-expressing cells. In vitro methylation at a cytosine residue within a footprint region prevents the interaction of a factor with its binding site.

Introduction

Transcription of genes by RNA polymerase II is brought about by sequence-specific DNA-binding proteins that not only help to establish a basal transcription rate but also confer specificity with regard to cell type, developmental timing, and environmental responsiveness to a given promoter (for review, see Dynan and Tjian, 1985; Maniatis et al., 1987). These binding proteins have been characterized in crude nuclear extracts (Dignam et al., 1983; Wildeman et al., 1984) by their ability to bind to specific DNA sites as determined by gel retardation (Fried and Crothers, 1981; Garner and Revzin, 1981) and DNAase I footprinting techniques (Galas and Schmitz, 1978). The power of these in vitro approaches has led to tremendous progress in defining *trans*-acting DNA-binding activities as well as their target elements within promoter and enhancer sequences. Given the highly organized chromatin structure into which DNA within a eukaryotic nucleus is compacted (for review, see Igo-Kemenes et al., 1982; Eissenberg et al., 1985), no conclusions on whether these binding sites are indeed occupied by proteins in the intact cell can be derived from in vitro studies only. The genomic sequencing methodology introduced by Church and Gilbert (1984) enabled for the first time the direct analysis of factors binding to their target sites in living cells. Its most powerful application so far has been the footprinting of proteins in vivo with dimethylsulfate (DMS), a reagent that readily enters the nucleus of an intact cell, thereby allowing the analysis of DMS reactivity of guanosines (Ogata and Gilbert, 1978) without prior nuclei isolation. The in vivo footprinting methodology has been applied successfully to prokaryotes (Nick and Gilbert, 1985; Martin et al., 1986), yeast (Giniger et al., 1985), and mammals (Ephrussi et al., 1985; Becker et al., 1986; Pauli et al., 1987).

To date, the only case where DNA-binding activities in extracts have been related to actual in vivo interactions of factors with DNA is that of studies on the immunoglobulin heavy chain gene enhancer. Analyses of protein contacts within a stretch of genomic DNA of known enhancer activity (performed by Ephrussi et al., 1985, who used the genomic footprinting procedure) revealed five distinct protein-binding sites (E1–E4 and O; Weinberger et al., 1986) that were found in B cells but not in T cells or erythroid cells. Interestingly, protein factors binding to sites E1, E3, and O could be assayed not only in extracts from B cell nuclei but also in those of T cells, erythroid cells, and fibroblasts (Weinberger et al., 1986; Sen and Baltimore, 1986; Schlokat et al., 1986). On the other hand, no in vitro footprint has yet been obtained of the in vivo site E2 (Augerau and Chambon, 1986). The in vitro analyses are further complicated by the existence of ubiquitous factors that share a target sequence with tissue-specific subclasses (Gerster et al., 1987; Landolfi et al., 1986; Staudt et al., 1986). It is not yet clear how general the implications of these findings are; however, they emphasize the importance of in vivo analyses in conjunction with standard in vitro approaches.

We are studying transcriptional regulation of the rat tyrosine aminotransferase (TAT) gene since it is subject to a variety of regulatory mechanisms. Following its developmentally controlled onset (Greengard, 1970) by a proposed *trans*-acting factor (for review, see Gluecksohn-Waelsch, 1987), the gene is transcribed in the parenchymal cells of liver (for review, see Hargrove and Granner, 1985) with a low but constant level. The transcription rate can be increased by either glucocorticoids or glucagon via its intracellular mediator, cAMP (Hashimoto et al., 1984; Schmid et al., 1987). The initial indications of sequences that might be important for transcriptional regulation were obtained from analyses of chromatin structure in different tissues. Specific sites in the 5′ flanking region of the TAT gene were found to be hypersensitive (HS) to DNAase I in isolated nuclei of TAT-expressing but not of nonexpressing cells (Becker et al., 1984). They occur within the first 200 nucleotides upstream of the TAT cap site as well as around −1000 (HS sites I and II, respectively, in Figure 1). A third DNAase I–HS site (III in Figure 1) at about 2.5 kb upstream of the TAT-coding body appears upon induction of the gene by glucocorticoids (Becker et al., 1984). Indeed, Jantzen et al. (1987) have shown that two functionally cooperating glucocorticoid response elements (GREs) are contained within this HS site. Purified gluco-

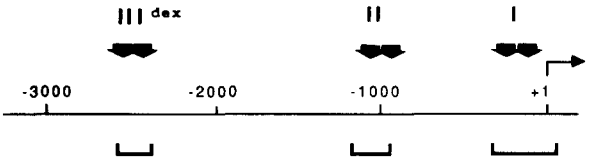

Figure 1. DNAase I–Hypersensitive Sites within the 5′ Flanking Region of the Rat TAT Gene

The first 3 kb of 5′ flanking sequence of the TAT gene are depicted. DNAase I–hypersensitive sites (Becker et al., 1984) are indicated by double arrows. The glucocorticoid-inducible HS site III is labeled "dex." Brackets mark the regions that have been analyzed by in vivo footprinting either in Becker et al. (1986) (HS site III) or in this study (HS sites I and II).

corticoid receptor binds to these GREs in vitro (Jantzen et al., 1987), and with genomic footprinting technology it was shown that glucocorticoid receptor as well as other factors interact with the GREs in vivo in a hormone-dependent manner (Becker et al., 1986).

The identification of the functional GREs of the TAT gene at a site that revealed itself by an alteration of chromatin structure prompted us to study the proteins binding to DNA within DNAase I–HS sites I and II (Figure 1) in the living cell as well as by standard in vitro approaches. The results presented here thus contribute to defining the relationship between sets of data that have been obtained either in vitro or in vivo. It is a compelling conclusion that specific protein–DNA interactions within cells are not determined merely by the presence of the protein factors, but also by the accessibility of their target sequences.

Results

Proteins Interact with DNA within DNAase I–HS Sites I and II In Vivo Only in TAT-Expressing Hepatoma Cells

While evidence for protein–DNA interactions derived from analyses of isolated nuclei with either DMS (Church et al., 1985) or DNAase I (Jackson and Felsenfeld, 1985; Zinn and Maniatis, 1986) has been reported, our earlier studies (P. B. Becker, unpublished data) showed that most of the reactivity changes at guanosines obtained from DMS treatment of whole cells could not be reproduced in nuclei that had been prepared by standard techniques. It was therefore assumed that nuclei are already artifactual to varying degrees because of experimental manipulation (most likely from leakage of factors), and no further attempt was made to footprint factors in nuclei with DNAase I.

In vivo footprints were obtained by suspension of living cells in culture medium and reaction with dimethylsulfate (DMS) under conditions that result in a partial methylation of N_7 residues of guanosines (Ephrussi et al., 1985). Close contact of a protein to the DNA can either decrease the reactivity of guanosines (Gs) by tight binding within the major groove or enhance their reactivity, presumably by locally increasing the reagent concentration in hydrophobic patches (Ogata and Gilbert, 1978). After in vivo methylation the genomic DNA is purified, cleaved with a

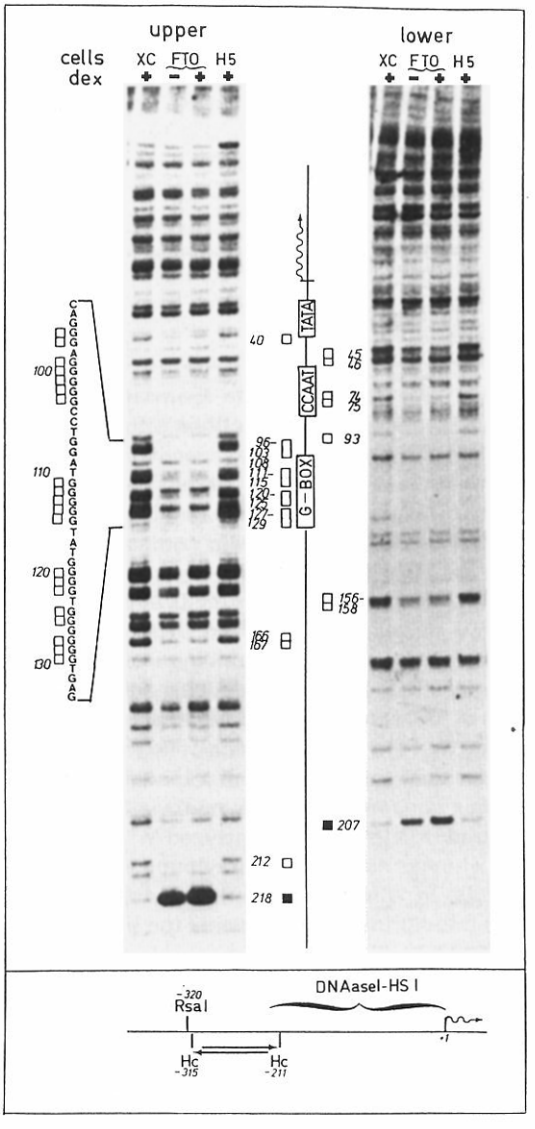

Figure 2. In Vivo DMS Reactivity Experiments Reveal Protein–DNA Interactions within HS Site I

XC, H5, and FTO-2B cells either induced with dexamethasone (+) or not (−) were incubated with DMS. After purification DNA was restricted with RsaI, and the DNA backbone was cleaved at positions of modified guanosines with piperidine. Thirty micrograms of total genomic DNA was run on a 6% denaturing polyacrylamide gel, transferred onto GeneScreen membrane, and covalently crosslinked by UV irradiation. The DNA was then hybridized with the single-stranded DNA probe fragment indicated in the lower part of the figure to probe for the upper strand (left). After autoradiography the membrane was treated with NaOH to remove the bound probe, and was rehybridized with the complementary probe fragment to visualize the lower strand (right). Guanosines that are altered in their reactivity in FTO-2B cells as compared to XC and H5 cells are marked as follows: enhancements with a filled square, protections with an open square. Along the central vertical line some sequences with homologies to well-known functional elements, the location of the clustered protections ("G-Box") as well as the approximate position of the transcriptional start site (wavy line), are indicated. The lower part of the figure shows the location of the DNAase I–hypersensitive site I in the immediate vicinity of the TAT cap site (+1), and the location of the probes used. Numbers indicate the positions of the nucleotides upstream of the major transcriptional start site (Shinomiya et al., 1984).

suitable restriction enzyme, and finally reacted with piperidine to break the DNA backbone at positions of guanosine modification (Maxam and Gilbert, 1980). Aliquots of the DNA are run on polyacrylamide sequencing gels, transferred to a nylon membrane, and covalently bound by UV crosslinking. The guanosine sequence ladder of the region of interest is visualized selectively among the rest of the genome by indirect end-labeling (Church and Gilbert, 1984).

Figure 2 shows the result of such an experiment, where the first 220 nucleotides upstream of the TAT cap site have been analyzed. Most instructive is a comparison of the reactivity of each guanosine residue (measured by the intensity of the corresponding band) among the following: the TAT-expressing rat hepatoma cell line FTO-2B (Killary et al., 1984), the nonexpressing dedifferentiated hepatoma variant H5 (Weiss et al., 1981), and rat fibroblast XC cells (Svoboda, 1960). Whereas the pattern obtained from H5 and XC cells is essentially like the one from partially methylated protein-free genomic DNA (not shown), multiple bands in the FTO-2B DNA appear to be protected or enhanced when compared to the ones from nonexpressing cells or control DNA. The changes in reactivity marked in Figure 2 occur in discrete patches, indicating the presence of several distinct protein-binding sites. The 5′ boundary of the recorded effects is marked by a strong enhancement in guanosine reactivity at G −218 on the upper and G −207 on the lower strand. Given the helical turn of the DNA, the enhancements G −218 and G −207 thus adjoin each other, and the protection of G −212 occurs between them on the opposite side of the helix. The most prominent protections within the region analyzed are found at clusters of guanosines on the upper strand marked "G-box" in Figure 2 (positions −96 to −130), which are reminiscent of a protein-binding site containing a string of 16 guanosine residues upstream of the β-globin gene (Jackson and Felsenfeld, 1985). The reactivity of most of the guanosines present in these clusters is influenced in FTO-2B cells. Several protected residues can be scored between the G-box and the cap site, notably protections of the Gs residing in CCAAT-box homologies at −75 and at −40, the latter being close to the presumed TATA box. A summary of the recorded effects is displayed in Figure 5B.

All guanosines influenced in their reactivity are contained within the region that is found to be hypersensitive to DNAase I in nuclei of TAT-expressing cells. No indications of further protein–DNA interactions were detected in an analysis of the neighboring DNA up to −350 or +50, using other probe fragments (data not shown). No alterations in DMS reactivity were observed after the transcription rate of the TAT gene was increased by induction with glucocorticoids (Figure 2).

Experiments similar to the one described for HS site I were carried out by analyzing the sequences between −1000 and −1100 that are found to be hypersensitive in nuclei of FTO-2B cells (HS site II). Figure 3 displays guanosine reactivities of DNA from FTO-2B cells compared to those obtained from H5 and XC cells, as well as in vitro–methylated protein-free genomic DNA ("Co").

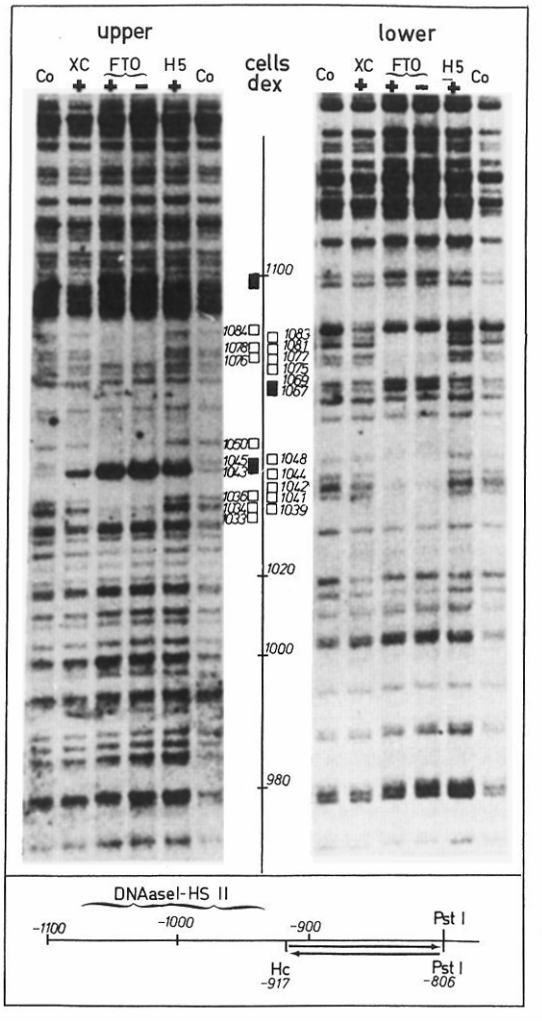

Figure 3. In Vivo DMS Reactivity Experiments Reveal Protein–DNA Interactions within HS Site II

After in vivo methylation, purified DNAs were restricted with PstI and further processed as described in Figure 2. The probes used to visualize the genomic sequences are depicted in the lower part of the figure. In this experiment, control reactions with protein-free genomic DNA ("Co") are also shown: 30 μg of protein-free genomic DNA was cleaved with PstI before being subjected to a standard DMS reaction and piperidine cleavage according to Maxam and Gilbert (1980). Symbols, abbreviations, and numbering are as detailed in Figure 2.

Striking effects again were found in the TAT-expressing FTO-2B cells. Two clusters of G protections with interspersed (G −1043, G −1045) or flanking (Gs −1067, −1069, −1098, −1099) enhancements define the binding sites for at least two protein factors between nucleotides −1030 and −1100 upstream of the transcription start site. No changes in reactivity pattern were observed after induction of the gene with dexamethasone. Aliquots of DNA from the same in vivo methylation reaction that show hormone-dependent changes at −2.5 kb (HS III; Becker et al., 1986) served as internal controls in these experiments. The results are summarized in Figure 5A.

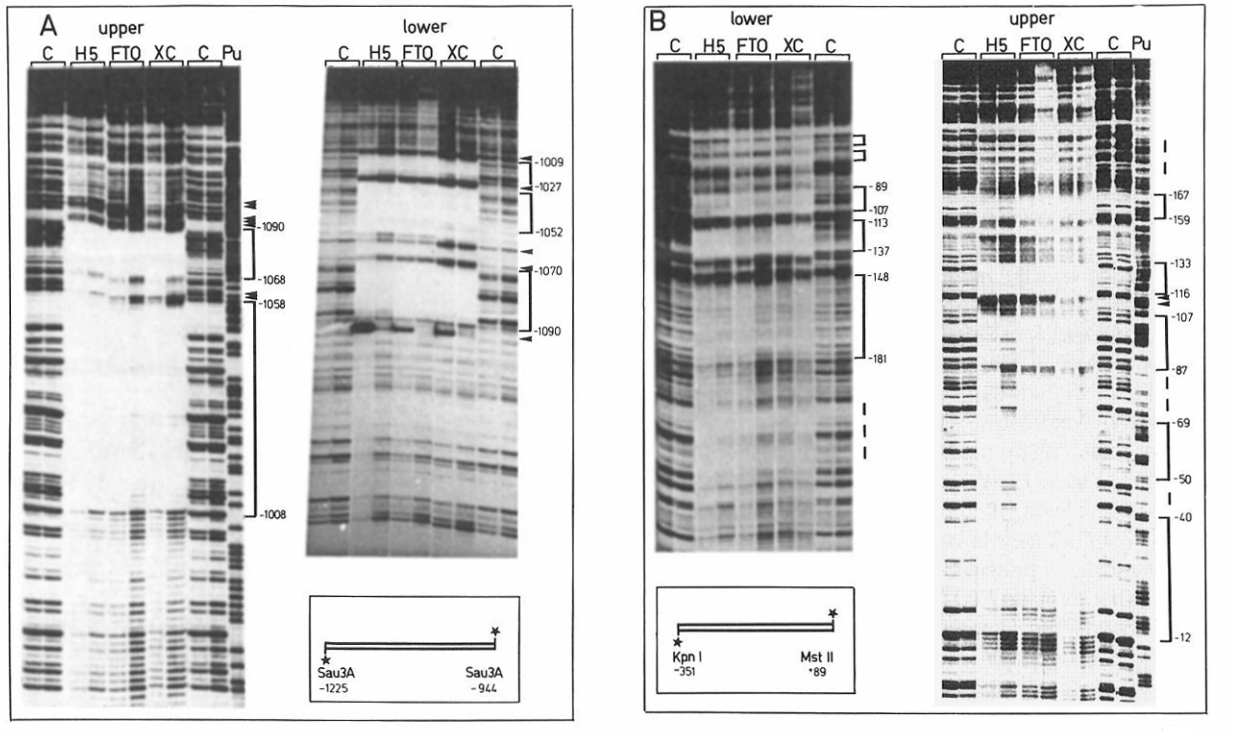

Figure 4. Nuclear Proteins from TAT-Expressing and Nonexpressing Cell Lines Bind to TAT Upstream DNA In Vitro

(A) Twenty-seven micrograms of extract protein from FTO-2B, H5, and XC cells was incubated with 100 ng of sonicated salmon sperm DNA and the indicated end-labeled fragment, as detailed in Experimental Procedures. DNAase I was added to a final concentration of 10 or 12 μg/ml and the digestion allowed to proceed for 90 sec at 20°C. After the reaction, purified DNA was separated on an 8% denaturing polyacrylamide gel. The gel was removed from the glass plate, dried, and exposed to x-ray film. For controls ("C"), fragments were digested with 1.5 and 2 μg/ml of DNAase I in the absence of extract protein. Brackets show regions protected from DNAase I digest in extract samples. Arrows indicate positions of enhanced DNAase I cleavages. A purine-specific ("Pu") sequence reaction indicates the positions of the corresponding nucleotides (numbers) upstream of the TAT cap site.

(B) The experiment was essentially as described in (A) except that the indicated fragment (KpnI–MstII) was used and 46.5 μg of extract protein was incubated with 300 ng of sheared salmon sperm DNA. DNA was digested with 16 or 18 μg/ml (for the lower strand) and 24 or 28 μg/ml of DNAase I (for the upper strand). For control digests without extracts, 3 or 4 μg/ml of DNAase I was used.

Cell Type–Specific DNA Binding Is an In Vivo Quality of Ubiquitous Factors

We next proceeded to correlate the alterations in G reactivities observed above with DNA-binding activities in crude nuclear extracts (Dignam et al., 1983; Wildeman et al., 1984) from FTO-2B, H5, and XC cells. Because methylation reactivity experiments with these unfractionated extracts did not yield satisfying results, we performed conventional footprinting experiments using DNAase I (Galas and Schmitz, 1978).

Figure 4A shows such an in vitro footprinting experiment using an end-labeled fragment covering the region of DNA between −944 and −1225 bp upstream of the TAT cap site for the upper and the lower strand. Stretches of DNA that are shielded from DNAase I attack by extract proteins, as compared to control digests where no extract was added, are obvious on both strands. Identical DNAase I footprints were obtained when extracts from FTO-2B, H5, and XC cells were used, indicating that nuclei of all three cell lines contain the same binding activities. This result contrasts with the in vivo analyses (Figure 3), which indicated that the corresponding binding sites are occupied only in FTO-2B cells. We conclude that the presence of factors in nuclei is not sufficient for their interaction with potential target sites in vivo.

Similar analyses were performed to visualize protein footprints on a fragment spanning the TAT cap site (from +81 to −351), a region that has revealed multiple binding sites in vivo (Figure 2). Figure 4B shows an array of protected regions that suggests a close spacing of several protein-binding sites on the analyzed fragment. Again the footprinted regions appear to be very similar when extracts of all three cell lines are tested, leaving the same apparent discrepancy between in vivo and in vitro binding as discussed above. Minor differences between the footprints obtained from the three extracts are most likely due to experimental variation. The present state of analysis, however, does not allow us to discriminate between differences in binding proteins in the various extracts. The results obtained in vitro and in vivo are compared and related to the underlying DNA sequence in Figure 5. Generally, guanosine residues identified as targets for

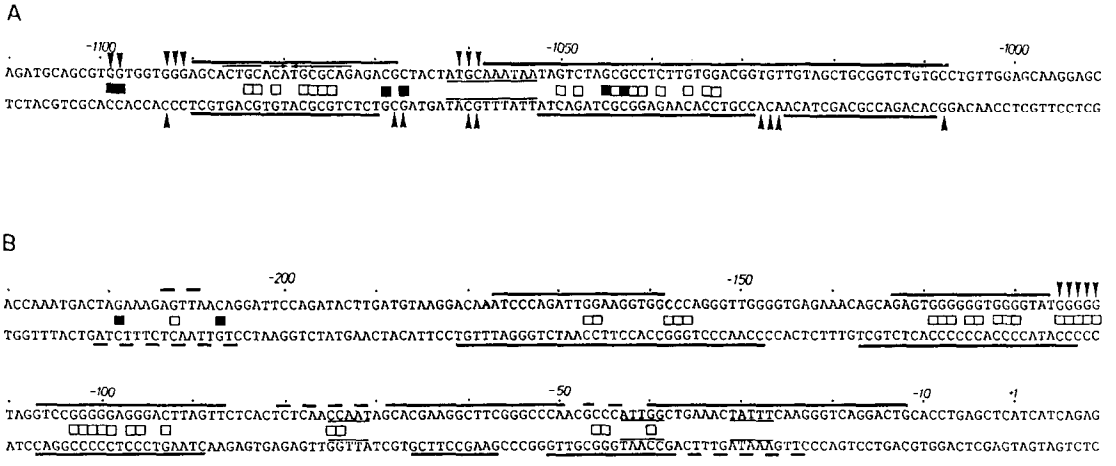

Figure 5. Summary of Protein–DNA Interactions In Vivo and In Vitro

The DNA sequences between positions −1110 and −990 (A) and between −230 and the cap site (+1) (B) are shown. Homologies to known functional sequences are indicated with a thin line. Interestingly, a sequence homology to the decamer motif (Falkner et al., 1986) is found between −1052 and −1062, which is not, however, contacted by a protein in vivo or in vitro. Guanosines of altered reactivity toward DMS in vivo (Figures 2 and 3) are marked with open squares for protections and filled squares for enhancements. Solid lines indicate the areas protected from DNAase I digestion by extract proteins (Figure 4). Dashed lines indicate ambiguities in footprints either due to weak footprints or at sites that are not cleaved by DNAase I in control DNA. Arrows indicate positions of enhanced DNAase I cuts due to protein binding.

protein binding in vivo are contained within the DNAase I footprints. It has not been proven that the factors characterized here by in vitro DNAase I footprinting are identical to those indicated by DMS reactivity in intact cells. However, the generally good agreement of the binding sites derived from the two sets of data suggests that this might be the case. While the overall consistency between the two sets of data is pleasing, there are a few cases where they differ. The CCAAT-box homology at −75, for example, is protected well in vivo but only weakly in vitro. The most striking difference occurs at a cluster of five guanosines (G −111 to G −115) that are clearly reduced in their reactivity in vivo but appear to be targets for enhanced DNAase I cleavage in extracts. Furthermore, there are stretches of DNA protected from DNAase I cleavage in vitro that do not contain protected Gs in vivo (e.g., −69/−50 or −40/−12). Whether these footprint extensions represent binding sites of factors not found in in vivo DMS reactivity analyses must be clarified by assaying protein binding on specifically mutagenized DNA fragments.

Hypomethylation In Vivo Correlates with Active TAT Gene Transcription

As shown above, many proteins that bind to specific target sequences in a cell type–specific manner are available in nuclei of cells where no such in vivo DNA interaction can be observed. Chromatin features such as specific nucleosome positioning, the folding of the DNA into a higher-order structure, or a stable modification of the DNA itself could account for this exclusion of binding proteins. A common modification of DNA is methylation of cytosines in CpG dinucleotides at their C_5 position (for review, see Bird, 1986), and numerous studies have established a certain correlation between hypomethylation and gene activity (for review, see Doerfler, 1983).

The genomic sequencing methodology employing the different reactivities of cytosine and 5-methylcytosine to hydrazine allows the analysis of the methylation status of CpG dinucleotides within the mammalian genome irrespective of whether these sites are contained within methylation-sensitive restriction sites (Church and Gilbert, 1984; Nick et al., 1986; Saluz et al., 1986). The presence of clustered CpG dinucleotides within HS sites I and II (Figure 6) prompted us to analyze the methylation status of these sites in the genome of the three cell lines utilized. The experiment (Figure 6) shows that bands corresponding to all cytosines contained in CpG dinucleotides are present in FTO-2B cells at intensities comparable to other cytosines. Therefore, all methylation targets are unmethylated in FTO-2B cells. In the nonexpressing cells H5 and XC, all CpG sequences between −1000 and −1150 are completely methylated on both strands, as the corresponding bands are missing. This is also true for the three well-resolved CpG sites (C −47, C −56, C −65) in H5 cells, whereas XC DNA appears to be heterogeneous with regard to these latter sites: the indicative bands are clearly visible, though with reduced intensities. The study has been extended by analysis of the methylation status of all previously mentioned CpG sites in some rat tissues (not shown). Whereas all sites are unmethylated in rat liver, they are completely modified in brain, kidney, and spleen. Thus a good correlation between the absence of methylation and TAT gene transcription is observed, while in general all analyzed CpGs are heavily methylated in nonexpressing cells.

Binding of a Protein to DNA Is Prevented by Cytosine Methylation

The fortuitous finding that the two CpG dinucleotides in the center of each of the two footprints between −1000

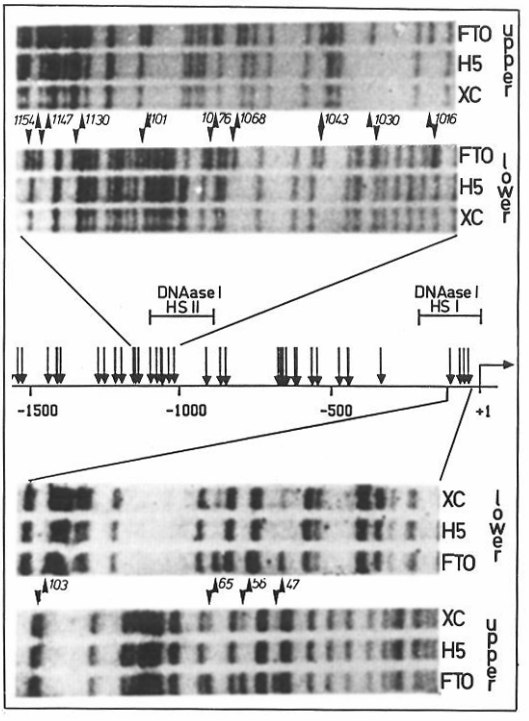

Figure 6. CpG Methylation at Sites of Protein–DNA Interaction in Genomic DNA

The central horizontal line depicts a stretch of 1.5 kb of 5′ flanking DNA to the TAT transcriptional start site (+1). Each vertical arrow marks the position of a genomic CpG dinucleotide. The regions between −1000 and −1160 (upper part) and between −10 and −110 (lower part) were analyzed for cytosine methylation by genomic sequencing. Thirty micrograms of protein-free DNA from FTO-2B, H5, and XC cells was restricted with the appropriate restriction enzymes and subjected to a standard cytosine-specific reaction with hydrazine (Maxam and Gilbert, 1980). After piperidine treatment, the genomic fragments were separated on a denaturing polyacrylamide gel and covalently bound to GeneScreen membrane by UV irradiation. Both strands were visualized by successive hybridizations of the same membrane with complementary single-stranded DNA probes indicated in Figure 2 (lower part) and Figure 3 (upper part). Arrowheads mark bands corresponding to cytosines contained in CpG dinucleotides.

and −1100 (C −1076 and C −1043 on the upper strand) are contained in HhaI restriction sites allowed the direct assay of whether cytosine methylation affects protein binding to DNA. The end-labeled fragment used for footprinting was completely methylated at the two sites with HhaI methylase. Control fragments were processed in parallel without the addition of methylase ("mock"). Both labeled DNAs were then used for footprinting with extracts from FTO-2B and H5 nuclei (Figure 7). On the mock-methylated fragments the expected footprints (compare to Figure 4A) are readily obtained, whereas methylation severely influences protein binding. The distal footprint is completely abolished, and an additional DNAase I cleavage appears within the proximal one at the position of the methylated CpG residue. Both FTO-2B and H5 extracts give identical results in this experiment.

Discussion

Cell type–specific interaction of factors with their target sites in responsive genes may be controlled by the mere presence of factors in TAT-expressing cells, by changes in affinity for their binding sites, or by modulation of the proteins' access to their sites of action. The presence of cell type–specific factors in nuclear extracts from various cells has recently been described in binding (Gerster et al., 1987; Landolfi et al., 1986; Staudt et al., 1986) as well as in in vitro transcription assays of tissue-specific promoters (Gorski et al., 1986; Bodner and Karin, 1987). As all DNA-binding activities described to interact with the 5′ region of the TAT gene are present in extracts of cells in which no specific protein–DNA interactions can be detected in vivo, the mere presence or absence of these factors is not sufficient to explain why their binding sites are not contacted in nonexpressing cells. The experiments presented here do not, however, rule out the possibility that the identified sequence elements are contacted by different proteins present in the three extracts that produce identical footprints when assayed with DNAase I. Such a case has been reported for the octamer sequence of the immunoglobulin κ gene promoter (Staudt et al., 1986).

Induced protein binding after increasing the affinity for its binding site by protein modification is most likely a common mechanism involved in transcriptional control (Zimarino and Wu, 1987; Angel et al., 1987; Lee et al., 1987; for review, see Maniatis et al., 1987). Furthermore, the affinity of a binding protein can be modulated by its association with a specific ligand, as is the case for the steroid-induced specific DNA binding of steroid receptor (Becker et al., 1986; for review, see Yamamoto, 1985). While a quantification of the binding affinity of a factor in unfractionated extracts is difficult, results not shown suggest that the affinities of the described TAT factors in the various extracts do not differ greatly. Extract preparations from nuclei of all three cell lines yield comparable amounts of protein per number of cells. Equal amounts of protein were used to obtain footprints, and the optimal binding conditions (such as the amount of unspecific competitor DNA included in the binding) were identical for all three extracts. The two footprints at −1000 are competed for by similar amounts of unspecific or specific unlabeled competitor, and were equally stable over a period of more than 15 min at room temperature when challenged with a 100-fold excess of specific competitor fragment after binding (not shown). Binding activities in H5 and FTO-2B nuclear extracts are similarly influenced by cytosine methylation of their target sites (Figure 7).

In the absence of any detectable difference in binding affinity in vitro, we favor the idea that changes in chromatin structure and/or DNA modification determine the binding of factors to their DNA targets. This level of regulation could involve the packaging of DNA into higher-order chromatin structures, or the presence or absence of a nucleosome at a protein-binding site. The proteins analyzed bind to DNA only in cells where the corresponding target sequences are located within hypersensitive chromatin

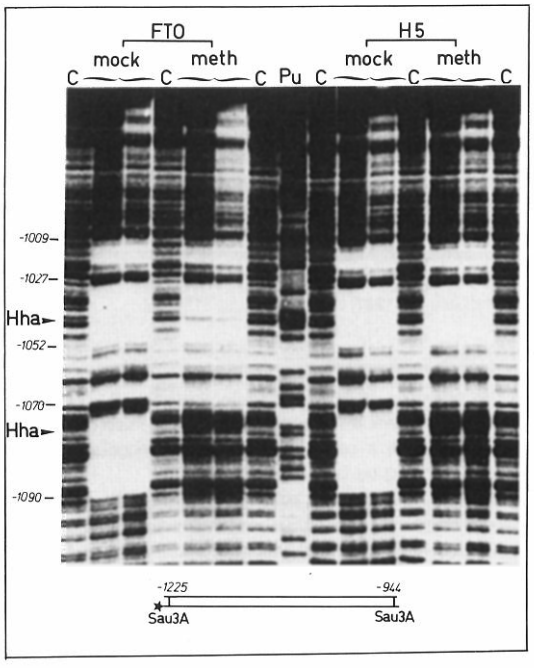

Figure 7. Cytosine Methylation Prevents Protein Binding

The end-labeled fragment indicated below the figure was treated with HhaI methylase to modify completely the two internal HhaI recognition sequences ("meth"). Control DNA was processed identically but without the addition of methylase ("mock"). Both DNAs were used in an in vitro footprinting experiment analogous to the one depicted in Figure 4A. Control digestions of fragment without extract ("C") were reacted with 2 μg/ml of DNAase I. "Pu" marks a purine-specific sequencing reaction. The positions of the methylated HhaI sites (G^mCGC) are indicated ("Hha"). Numbers refer to the positions of the corresponding nucleotides upstream of the TAT cap site.

structures. The causal relationship, however, between DNAase I hypersensitivity and protein binding to DNA remains unclear. In this context it will be of interest to compare the nucleosome positioning on the TAT gene in the different cell lines. That the presence of a nucleosome might interfere with binding of an available transcription factor has recently been suggested (Richard-Foy and Hager, 1987; Cordingley et al., 1987; Lorch et al., 1987). Stable modification of DNA in protein-binding sites may also prevent the interaction of ubiquitous factors in a nonexpressing cell. Data have been accumulated that ascribe methylation of cytosines in CpG dinucleotides a role in inactivating genes (for review, see Doerfler, 1983). More recently, results have been published that present evidence for a close link between the transcriptional onset of genes and site-specific demethylation both in long-term experiments involving stable transformation of cells with methylated templates (Yisraeli et al., 1986) and as a relatively short-term response upon estradiol stimulation of the chicken vitellogenin gene transcription (Saluz et al., 1986). All CpG dinucleotides within the regions of protein–DNA interactions of the TAT gene are heavily methylated in cell lines and tissues that do not express TAT, whereas they are not modified in FTO-2B cells and rat liver. Taking advantage of two HhaI methylation sites within protein-binding sites, we were able to show that methylation of a cytosine residue indeed strongly affects protein binding in vitro. This is not too surprising, as the C_5 position of cytosine reaches into the major groove like the N_7 of guanosine. It is still not clear whether it is methylation alone that prevents the factors from binding in vivo. This question, however, will be investigated by monitoring in vivo protein binding at these sites in nonexpressing cells after 5-azacytidine-induced demethylation.

Further support for the notion that chromatin structure and/or DNA modification affects gene expression comes from studies in which gene activity has been related to chromosomal position. Examples are overt position effects observed following stable integration of the rosy gene into various chromosomal loci, including the X chromosome (Spradling and Rubin, 1983), and the inactivation of a β-globin gene by a translocation event close by (Kioussis et al., 1983). In the latter case, DNAase I resistance and DNA methylation where shown to be correlated with gene inactivation.

The unexpected finding of cell type–specific DNA binding of ubiquitous proteins emphasizes the necessity of in vivo analyses. It also indicates levels of regulation of gene expression that are not sufficiently explained by presence and affinity of factors and that most likely involve changes in chromatin structure and/or DNA modification.

Experimental Procedures

Cell Culture

FTO-2B, H5, and XC cells were grown in DMEM with 10% fetal calf serum, 2 mM glutamine, 100 U/ml penicillin–100 μg/ml streptomycin, and 10 mM HEPES (pH 7.4) to confluency. The day before harvesting, the cells were incubated with serum-free DMEM for 16–20 hr followed by a change to medium that did or did not contain 10^{-6} M dexamethasone. Induction times ranged from 1–4 hr.

In Vivo Methylation of Cells and Preparation of DNA

Routinely, 1–2 × 10^8 cells were trypsinized mildly and resuspended in 1 ml of culture medium in the presence or absence of dexamethasone. The methylation with dimethylsulfate followed the protocol detailed by Ephrussi et al. (1985). After methylation, nuclei were prepared (Becker et al., 1984) and finally resuspended in 1 ml of 0.5 M EDTA. Sarcosyl and RNAase A (Sigma) were added to 0.5% and 200 μg/ml, respectively, and the reaction was incubated at 37°C for 3 hr. Proteinase K was added to 200 μg/ml, followed by an overnight incubation at 37°C. DNA was extracted with phenol and phenol/chloroform (1:1), dialyzed overnight against 3 l of 10 mM Tris, 1 mM EDTA, and finally precipitated. Control reactions with protein-free DNA as well as hydrazine reactions of genomic DNA were according to the sequencing protocol of Maxam and Gilbert (1980).

Preparation of Membranes for Hybridization

Thirty micrograms of genomic DNA was digested with 2 U/μg of restriction enzyme overnight. Digestion was stopped by adjusting to 10 mM EDTA; DNA was recovered by precipitation and subjected to piperidine treatment (Maxam and Gilbert, 1980). After the reaction, the solution was transferred into a fresh tube and DNA was precipitated, washed with 80% ethanol, and dried for several hr in the SpeedVac concentrator. DNA was washed and dried twice with 100 μl of water and finally dissolved in 3 μl of formamide loading buffer. Separation of the DNA on a 6% denaturing polyacrylamide gel and electroblotting onto a

GeneScreen membrane were as described by Church and Gilbert (1984). After the transfer was completed, the membrane was air-dried and baked for 20 min at 80°C in vacuo followed by 20 sec UV irradiation of the DNA-binding side from six germicidal tubes at 20 cm distance (5000 $\mu W/cm^2$) (Becker and Schütz, 1988).

Probe Fragments and Probe Synthesis

All probe fragments were cloned into Sma-cut M13mp8 (Messing and Vieira, 1985), and their identity and orientation were verified by sequencing (Sanger et al., 1977). Probe fragments used for genomic blots are indicated below the relevant figures. Probe syntheses from single-stranded M13 templates were as described by Church and Gilbert (1984) with modifications (Becker et al., 1986). A detailed description is published elsewhere (Becker and Schütz, 1988).

Hybridizations

Hybridizations and washes of membranes were performed as detailed by Church and Gilbert (1984) with the following modifications (Becker and Schütz, 1988). After the single-stranded probe DNA was recovered by isotachophoresis (Öfverstedt et al., 1984), it was incubated with a 100-fold molar excess of complementary sheared single-stranded M13 vector DNA in 5× SSC at 65°C for 1 hr to saturate vector-specific sequences. Hybridizations were carried out in polypropylene cylinders rotating in an incubator purchased from Bachhofer GmbH, Reutlingen, FRG. An extra 100 mM of NaCl was added to the washes per 5% of reduced GC contents (from 50%). For rehybridizations, membranes were washed for 15 min in 1 l of 200 mM NaOH at room temperature followed by two washes of 75 mM Na_2HPO_4 adjusted to pH 7.2.

Nuclear Extracts and DNAase I Footprinting

Preparation of extracts from nuclei were essentially as described by Dignam et al. (1983) and modified by Wildeman et al. (1984). Some extracts were finally dialyzed against 20 mM HEPES (pH 7.9), 20% glycerol, 100 mM KCl, 0.2 mM EDTA. The DNA fragments used for in vitro footprinting, a 432 bp long MstII (+81)/KpnI(−351) fragment and a 281 bp long Sau3A (−944)/Sua3A (−1225) fragment, were cloned into pUC vectors. After linearization with either EcoRI or HindIII within the polylinker, the resulting 5′ overhangs were filled in with Klenow enzyme, labeling with [α-^{32}P]dATP (Maniatis et al., 1982). After the reaction, the labeled DNA was restricted with HindIII or EcoRI, respectively, and the fragment of interest was purified by agarose gel electrophoresis and isotachophoresis (Öfverstedt et al., 1984). For DNAase I footprints with the 281 bp Sau3A fragment, 27 μg of extract protein was incubated with 100 ng of sheared salmon sperm DNA in a buffer consisting of 70 mM KCl, 10 mM HEPES (pH 7.9), 4 mM $MgCl_2$, 4 mM spermidine, 10% glycerol, 0.1 mM EDTA, in a total volume of 10 μl for 15 min on ice. To this, 1 ng of end-labeled DNA fragment in 1 μl of 10 mM Tris (pH 7.5), 0.1 mM EDTA, was added. The binding was allowed to take place for 15 min at 20°C. Two microliters of DNAase I (Cooper Biomedical, stored at 1 mg/ml in 150 mM NaCl, 50% glycerol, at −20°C) in buffer D (Dignam et al., 1983) was added to a final concentration of 8–12 μg/ml and again incubated at 20°C for 90 sec. To analyze footprints on the fragment containing the TAT cap site, 46.5 μg of extract protein, 300 ng of salmon sperm DNA, and 16–30 μg/ml of DNAase I were used. The reaction was stopped by addition of 120 μl of 0.3 M sodium acetate (pH 9.0), 10 mM EDTA, 200 μg/ml yeast total RNA, 0.1% SDS, 150 μg/ml Proteinase K. After 30 min of incubation at 50°C, the mixture was extracted with phenol/chloroform (1:1) and chloroform/isoamylalcohol (24:1), and was precipitated. Dry pellets were dissolved in 3 μl of formamide loading buffer, and DNA was separated on 6% or 8% denaturing polyacrylamide gels. Minor modifications of the footprinting procedure, such as using poly(dI-dC) as carrier DNA or reducing the salt to 30 mM KCl, did not change the results.

Methylation of Templates for In Vitro Footprinting

End-labeled fragments were methylated with HhaI methylase (Boehringer) according to the supplier's specifications. Complete methylation was checked by restriction analysis. Mock-methylated control DNA was processed in an identical way, except that no methylase was added.

Acknowledgments

We wish to thank M. Frick and Dr. D. Bohmann for an introduction to extract preparation; Dr. A. F. Stewart for many helpful discussions and important suggestions; Drs. M. Boshart, A. Nordheim, and A. F. Stewart for critical reading of the manuscript; W. Fleischer for photography; and P. Di Noi for outstanding secretarial assistance.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *"advertisement"* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received July 16, 1987; revised August 31, 1987.

References

Angel, P., Imagawa, M., Chiu, R., Stein, B., Imbra, R. J., Rahmsdorf, H. J., Jonat, C., Herrlich, P., and Karin, M. (1987). Phorbol ester-inducible genes contain a common *cis* element recognized by a TPA-modulated *trans*-acting factor. Cell *49*, 729–739.

Augerau, P., and Chambon, P. (1986). The mouse immunoglobulin heavy-chain enhancer: effect on transcription in vitro and binding of proteins present in HeLa and lymphoid B-cell extracts. EMBO J. *5*, 1791–1797.

Becker, P. B., and Schütz, G. (1988). Genomic footprinting. In Genetic Engineering, Principles and Methods, Vol. 10, J. K. Setlow, ed. (New York: Plenum Press), in press.

Becker, P. B., Renkawitz, R., and Schütz, G. (1984). Tissue-specific DNase I-hypersensitive sites in the 5′-flanking sequences of the tryptophan oxygenase and the tyrosine aminotransferase genes. EMBO J. *3*, 2015–2020.

Becker, P. B., Gloss, B., Schmid, W., Strähle, U., and Schütz, G. (1986). In vivo protein/DNA-interactions within a glucocorticoid response element requires presence of the hormone. Nature *324*, 686–688.

Bird, A. P. (1986). CpG rich islands and the function of DNA methylation. Nature *321*, 209–213.

Bodner, M., and Karin, M. (1987). A pituitary-specific *trans*-acting factor can stimulate transcription from the growth hormone promoter in extracts of nonexpressing cells. Cell *50*, 267–275.

Church, G. M., and Gilbert, W. (1984). Genomic sequencing. Proc. Natl. Acad. Sci. USA *81*, 1991–1995.

Church, G. M., Ephrussi, A., Gilbert, W., and Tonegawa, S. (1985). Cell-type-specific contacts to immunoglobulin enhancers in nuclei. Nature *313*, 798–801.

Cordingley, M. G., Riegel, A. T., and Hager, G. L. (1987). Steroid-dependent interaction of transcription factors with the inducible promoter of mouse mammary tumor virus in vivo. Cell *48*, 261–270.

Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolate mammalian nuclei. Nucl. Acids Res. *11*, 1475–1489.

Doerfler, W. (1983). DNA methylation and gene activity. Ann. Rev. Biochem. *52*, 93–124.

Dynan, W. S., and Tjian, R. (1985). Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. Nature *316*, 774–778.

Eissenberg, J. C., Cartwright, J. L., Thomas, G. H., and Elgin, S. C. (1985). Selected topics in chromatin structure. Ann. Rev. Genet. *19*, 485–536.

Ephrussi, A., Church, G. M., Tonegawa, S., and Gilbert, W. (1985). B lineage–specific interactions of an immunoglobulin enhancer with cellular factors in vivo. Science *227*, 134–140.

Falkner, F. G., Mocikat, R., and Zachau, H. G. (1986). Sequences closely related to an immunoglobin gene promoter/enhancer element occur also upstream of other eukaryotic and prokaryotic genes. Nucl. Acids Res. *14*, 8819–8827.

Fried, M. G., and Crothers, D. M. (1981). Equilibria and kinetics of Lac repressor operator interaction by polyacrylamide gel electrophoresis. Nucl. Acids Res. *9*, 6505–6525.

Galas, D. J., and Schmitz, A. (1978). DNAase footprinting, a simple method for the detection of protein–DNA binding specificity. Nucl. Acids Res. *5*, 3157–3170.

Garner, M. M., and Revzin, A. (1981). A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the E. coli lactose operon regulatory system. Nucl. Acids. Res. *9*, 3047–3060.

Gerster, T., Matthias, P., Thali, M., Jiricny, J., and Schaffner, W. (1987). Cell type–specificity elements of the immunoglobulin heavy chain gene enhancer. EMBO J. *6*, 1323–1330.

Giniger, E., Varnum, S. M., and Ptashne, M. (1985). Specific DNA binding of GAL4, a positive regulatory protein of yeast. Cell *40*, 767–774.

Gluecksohn-Waelsch, S. (1987). Regulatory genes in development. Trends Genet. *3*, 123–127.

Gorski, K., Carneiro, M., and Schibler, U. (1986). Tissue-specific in vitro transcription from the mouse albumin promoter. Cell *47*, 767–776.

Greengard, O. (1970). The developmental formation of enzymes in rat liver. In Mechanisms of Hormone Action, Vol. 1, G. Litwack, ed. (New York: Academic Press), pp. 53–85.

Hargrove, J. L., and Granner, D. K. (1985). Biosynthesis and intracellular processing of tyrosine aminotransferase. In Transaminases, P. Christen and P. E. Metzler, eds. (New York: John Wiley & Sons, Inc.), pp. 511–532.

Hashimoto, S., Schmid, W., and Schütz, G. (1984). Transcriptional activation of the rat liver TAT gene by cAMP. Proc. Natl. Acad. Sci. USA *81*, 6637–6647.

Igo-Kemenes, T., Hörz, W., and Zachau, H. G. (1982). Chromatin. Ann. Rev. Biochem. *51*, 89–121.

Jackson, P. D., and Felsenfeld, G. (1985). A method for mapping intranuclear protein–DNA interactions and its application to a nuclease-hypersensitive site. Proc. Natl. Acad. Sci. USA *82*, 2296–2300.

Jantzen, H.-M., Strähle, U., Gloss, B., Stewart, F., Schmid, W., Boshart, M., Miksicek, R., and Schütz, G. (1987). Cooperativity of glucocorticoid response elements located far upstream of the tyrosine aminotransferase gene. Cell *49*, 29–38.

Killary, A. M., Lugo, T. G., and Fournier, R. E. K. (1984). Isolation of thymidine kinase deficient rat hepatoma cells by selection with bromodeoxyuridine, Hoechst 33258 and visible light. Biochem. Genet. *22*, 201–213.

Kioussis, D., Vanin, E., DeLange, T., Flavell, R. A., and Grosveld, F. G. (1983). β-Globin gene inactivation by DNA translocation in γβ thalassaemia. Nature *306*, 662–666.

Landolfi, N. F., Capra, J. D., and Tucker, P. W. (1986). Interaction of cell-type-specific nuclear proteins with immunoglobulin V_H promoter region sequences. Nature *323*, 548–551.

Lee, W., Mitchell, P., and Tjian, R. (1987). Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. Cell *49*, 741–752.

Lorch, Y., LaPointe, J. W., and Kornberg, R. D. (1987). Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. Cell *49*, 203–210.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Maniatis, T., Goodbourn, S., and Fischer, J. A. (1987). Regulation of inducible and tissue-specific gene expression. Science *236*, 1237–1245.

Martin, K., Huo, L., and Schleif, R. F. (1986). The DNA loop model for ara repression: AraC protein occupies the proposed loop sites in vivo and repression-negative mutations lie in these same sites. Proc. Natl. Acad. Sci. USA *83*, 3654–3658.

Maxam, A. M., and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. Meth. Enzymol. *65*, 499–560.

Messing, J., and Vieira, J. (1985). A new pair of M13 vectors for selecting either strand of double digest restriction fragments. Gene *19*, 269–276.

Nick, H., and Gilbert, W. (1985). Detection in vivo of protein/DNA interactions within the lac operon of Escherichia coli. Nature *313*, 795–797.

Nick, H., Bowen, B., Ferl, R. J., and Gilbert, W. (1986). Detection of cytosine methylation in the maize alcohol dehydrogenase gene by genomic sequencing. Nature *319*, 243–246.

Öfverstedt, L. G., Hammarström, K., Balgobin, N., Hjerten, S., Petterson, U., and Chattopadhyaya, J. (1984). Rapid and quantitative recovery of DNA fragments from gels by displacement electrophoresis (isotachophoresis). Biochim. Biophys. Acta *782*, 120–126.

Ogata, R. T., and Gilbert, W. (1978). An aminoterminal fragment of lac repressor binds specifically to lac operator. Proc. Natl. Acad. Sci. USA *75*, 5851–5854.

Pauli, U., Chrysogelos, S., Stein, G., Stein, J., and Nick, H. (1987). Protein–DNA interactions in vivo upstream of a cell cycle–regulated human H4 histone gene. Science *236*, 1308–1311.

Richard-Foy, H., and Hager, G. L. (1987). Sequence-specific positioning of nucleosomes over the steroid-inducible MMTV promoter. EMBO J. *6*, 2321–2328.

Saluz, H. P., Jiricny, J., and Jost, J. P. (1986). Genomic sequencing reveals a positive correlation between the kinetics of strand-specific DNA demethylation of the overlapping estradiol/glucocorticoid-receptor binding sites and the rate of avian vitellogenin mRNA synthesis. Proc. Natl. Acad. Sci. USA *83*, 7167–7171.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA *74*, 5463–5467.

Schlokat, U., Bohmann, D., Schöler, H., and Gruss, P. (1986). Nuclear factors binding specific sequences within the immunoglobulin enhancer interact differentially with other enhancer elements. EMBO J. *5*, 3251–3258.

Schmid, E., Schmid, W., Mayer, D., Jastorff, B., and Schütz, G. (1987). Transcriptional activation of the tyrosine aminotransferase gene by glucocorticoids and cAMP in primary hepatocytes. Eur. J. Biochem. *165*, 499–506.

Sen, R., and Baltimore, D. (1986). Multiple nuclear factors interact with the immunoglobulin enhancer sequences. Cell *46*, 705–716.

Shinomiya, T., Scherer, G., Schmid, W., Zentgraf, H., and Schütz, G. (1984). Isolation and characterization of the rat tyrosine aminotransferase gene. Proc. Natl. Acad. Sci. USA *81*, 1346–1350.

Spradling, A. C., and Rubin, G. M. (1983). The effect of chromosomal position on the expression of the Drosophila xanthine dehydrogenase gene. Cell *34*, 47–57.

Staudt, L. M., Singh, H., Sen, R., Wirth, T., Sharp, P. A., and Baltimore, D. (1986). A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. Nature *323*, 640–643.

Svoboda, J. (1960). Presence of chicken tumor virus in the sarcoma of the adult rat inoculated after birth with Rous sarcoma virus. Nature *186*, 980–981.

Weinberger, J., Baltimore, D., and Sharp, P. A. (1986). Distinct factors bind to apparently homologous sequences in the immunoglobulin heavy-chain enhancer. Nature *322*, 846–848.

Weiss, M. C., Sperling, L., Moore, E. E., and Deschatrette, J. (1981). Genetic and biochemical analysis of de-differentiated variants of a rat hepatoma. In Genes, Chromosomes and Neoplasia, Arrighi, Rao, and Stubblefield, eds. (New York: Raven Press), pp. 175–191.

Wildeman, A. G., Sassone-Corsi, P., Grundström, T., Zenke, M., and Chambon, P. (1984). Stimulation of in vitro transcription from the SV40 early promoter by the enhancer involves a specific *trans*-acting factor. EMBO J. *3*, 3129–3133.

Yamamoto, K. R. (1985). Steroid receptor regulated transcription of specific genes at gene networks. Ann. Rev. Genet. *19*, 209–252.

Yisraeli, J., Adelstein, R. S., Melloul, D., Nudel, U., Yaffe, D., and Cedar, H. (1986). Muscle-specific activation of a methylated chimeric actin gene. Cell *46*, 409–416.

Zimarino, V., and Wu, C. (1987). Induction of sequence-specific binding of Drosophila heatshock activator protein without protein synthesis. Nature *327*, 727–730.

Zinn, K., and Maniatis, T. (1986). Detection of factors that interact with the human β-interferon regulatory region in vivo by DNAase I footprinting. Cell *45*, 611–618.